Chapter 15: Are there differences between conditions or groups? Wilcoxon and Mann-Whitney U.

Full answers to study questions

1. Yes, the answers to all three of the questions are exactly the same!
   1.1. 7
   1.2. 7
   1.3. 7
2. The significance for each analysis would be:
   2.1. Significant with $p < .020$.
   2.2. Not significant.
   2.3. Significant with $p < .002$.
3. In a full write up you should also use descriptive statistics to interpret the direction of the difference. I've added these in, using the median, to show how an interpretation could be written up. Your exact wording could be quite different, and that is fine.
   3.1. The Mann-Whitney U showed a significant difference in addictive personality scores between addicts and non-addicts ($U = 9$, $p < .050$), with higher addiction scores for addicts (median = 85) than for non-addicts (median = 62).
   3.2. Addiction scores were higher before the treatment (median = 88) than after the treatment (median = 72), and a Wilcoxon signed rank test revealed that this was a significant reduction ($W = 16$, $p < .002$).

Full answers for additional dataset (Mann-Whitney U)

1. What method of analysis will you use to analyse this dataset?

   An experimental design was used, where the IV is whether the scent of chocolate is in the air while people shop. There are two independent conditions: no scent or chocolate scent. The DV is how much people spend. Due to an outlier in the dataset a non-parametric test was deemed necessary, and therefore a Mann-Whitney U test was conducted.
2. Suggest a suitable hypothesis for this analysis.

   There will be a difference in spending between the two conditions. (Note, this is a two tailed hypothesis. Either one or two tailed would be fine.)
3. Calculate the appropriate statistic and determine whether it is significant.

	No scent condition (Group 1)		Chocolate scent condition (Group 2)	
	Raw score	Rank	Raw score	Rank
	45	10.5	67	16
	36	5.5	45	10.5
	27	2	152	18
	56	12.5	59	14
	29	3	74	17
	36	5.5	56	12.5
	32	4	39	8
	41	9	63	15
	38	7		
	25	1		
Rank total		$T^1 = 60$		$T^2 = 111$

Calculating U_1 for the no scent condition

$$U_1 = (10 * 8) + \left(10 * \frac{(10 + 1)}{2}\right) - 60$$

Calculate the addition sum within the brackets.

$$U_1 = (10 * 8) + \left(10 * \frac{(11)}{2}\right) - 60$$

Calculate the division sum within the brackets.

$$U_1 = (10 * 8) + (10 * 5.5) - 60$$

Calculate the each of the multiplications within the brackets.

$$U_1 = (80) + (55) - 60$$

Finish the calculations by completing the addition and then the subtraction.

$$U_1 = 75$$

Calculating U_2 for the chocolate scent condition

$$U_2 = (10 * 8) + \left(8 * \frac{(8 + 1)}{2}\right) - 111$$

Calculate the addition sum within the brackets.

$$U_2 = (10 * 8) + \left(8 * \frac{(9)}{2}\right) - 111$$

Calculate the division sum within the brackets.

$$U_2 = (10 * 8) + (8 * 4.5) - 111$$

Calculate the each of the multiplications within the brackets.

$$U_2 = (80) + (36) - 111$$

Finish the calculations by completing the addition and then the subtraction.

$$U_2 = 5$$

4. Interpret and write up your findings using APA standards.

   A Mann-Whitney U analysis showed that the two groups differed significantly (U = 5, p < .050), with a median spend of £36 in the no scent condition and a median spend of £61 in the scent condition. (Note, this analysis assumes a two-tailed hypothesis).

5. Suggest three possible confounding variables for this study, and how you might measure each of these.

   There are many possible confounding variables, but here are three options I came up with.

   5.1. How much the participant likes chocolate. If this differed between the groups, then this might explain some of the difference.

   5.2. Participant's disposable income. If people have a higher disposable income, then they are likely to spend more. So if disposable income differed between the groups that could have a big influence on the findings. Imagine that those with higher income happen to be in the "chocolate scent" condition, then it might be that higher disposable income explains their higher level of spending, rather than the chocolate scent.

   5.3. Frequency of shopping. If someone shops frequently, then they may spend less per trip. Again, if this differs between the two conditions, it may change the magnitude of the difference between them.

Full answers for additional dataset (Wilcoxon signed rank test)

1. What method of analysis will you use to analyse this dataset?

   The study was an experiment with a repeated measures design. The IV was whether testing occurred before or after participating in a neurorehabilitation programme. The DV is memory score. Since there are some outliers in the form of very low scores, a non-parametric test will be used for the analysis, specifically Wilcoxon signed rank test.

2. Suggest a suitable hypothesis for this analysis.

   It is predicted that memory scores will increase after participating in a neurorehabilitation programme. (Note, this is a one tailed hypothesis. Either one or two tailed would be fine.)

3. Calculate the appropriate statistic and determine whether it is significant.

Before (Condition X)	After (Condition Y)	Difference (Y – X)	Rank	Signed rank	Sum of positive ranks	Sum of negative ranks
18	24	6	7.5	7.5	7.5	
12	12	0				
4	13	9	10.5	10.5	10.5	
15	14	-1	1.5	-1.5		-1.5
20	22	2	3.5	3.5	3.5	
2	9	7	9	9	9	
15	18	3	5.5	5.5	5.5	
13	22	9	10.5	10.5	10.5	
10	16	6	7.5	7.5	7.5	
21	24	3	5.5	5.5	5.5	
19	17	-2	3.5	-3.5		-3.5
17	18	1	1.5	1.5	1.5	
					$W+ = 61$	$W- = -5$

   The smaller of the *W* values is the calculated value, so $W = -5$.

4. Interpret and write up your findings using APA standards.

   Analysis using a Wilcoxon Signed Rank Test showed a significant improvement in memory ($W = -5$, $p < .005$), as scores were lower before the neurorehabilitation (median = 15) than afterwards (median = 17.5). (Note, this analysis assumes a one-tailed hypothesis).

5. The researchers used the same version of the memory test at both time points. Do you think this could be a methodological weakness? If so, why and how could you improve their design?

Using the same test risks practice effects, and therefore the improvement in memory may result from practice rather than the neurorehabilitation. To avoid this, the researchers could use two alternative versions of the same memory test.